bio-protocol Preprint

Not Peer Reviewed
Ask a question online

Computational prediction of RNA structures

Yan Zhang

Updated date: Nov 19, 2020

An abbreviated version of this protocol was published in eLIFE in Jan 2017

Operon mRNAs are organized into ORF-centric structures that predict translation efficiency

DOI: 10.7554/eLife.22037

Detailed protocol

Computational prediction of RNA structures using DMS-seq data

1. Raw DMS-seq data from different biological replicates were pooled after checking the reproducibility of the data (if applicable).
2. For the RNA region of interest, normalize the raw DMS-seq data to the most highly reactive residue within the region after removing outliers by 95% Winsorisation. In this procedure, all data above the 95th percentile is set to the 95th percentile.
3. A/C bases with a DMS-seq signal greater that 20% of the signal on the most highly reactive residue (after the normalization described in 2) were called 'unpaired'. The signal cutoff can be further optimized by using the receiver operating characteristic (ROC) curve that compares the DMS-seq signal of the RNA to the structure model obtained by other approaches (e.g. published crystal structure, if available).
4. Generate the RNA sequence of the region of interest.
5. Generate the folding constraint, by assigning the experimentally determined unpaired A/C bases as "x" and the bases with no constraint as ".".
6. For prediction of mRNA structures, use the ViennaRNA package (https://www.tbi.univie.ac.at/RNA/) or RNAfold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). This allows generation of a minimum free energy model of the indicated region, based on its RNA sequence (described in 4) and folding constraint (described in 5).
7. The mRNA structural model was visualized using VARNA (http://varna.lri.fr/), by importing the RNA sequence and the structure prediction in dot-bracket notation where the characters "(" and ")" correspond to the 5' base and the 3' base in the base-pair, respectively, while "." denotes an unpaired base.

How to cite:(Readers should cite both the Bio-protocol preprint and the original research article where this protocol was used)

1. Zhang, Y. (2020). Computational prediction of RNA structures. Bio-protocol Preprint. bio-protocol.org/prep641.
2. Burkhardt, D. H., Rouskin, S., Zhang, Y., Li, G., Weissman, J. S. and Gross, C. A.(2017). Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. eLIFE. DOI: 10.7554/eLife.22037